Executive summary

Extensive data are available on the concentrations of a wide range of chemicals in drinking water and also on the amount of water consumed by the public. The approach used to set drinking water quality standards and to do risk assessments only uses part of this data. Typically the highest measured concentration is combined with a high end consumption value to give a deterministic estimate of exposure. This can then be compared with a toxicological based intake that is considered to pose no significant risk over a lifetime's exposure. Such an approach affords a high degree of protection to consumers. However, it may be useful to know what the statistical distribution of exposure is likely to be, in order to understand the range of exposure within the population and to make better estimates of extreme exposure percentiles. Probabilistic assessments use the whole distribution for each variable to derive a distribution for the intake of the chemical from which summary statistics may be derived, if required.

The main objective of this project was to develop and explore the potential benefits of probabilistic approaches to exposure assessment to chemicals by ingestion of tap water in England and Wales.

Several approaches to probabilistic modelling were reviewed to assess their suitability for this purpose: analytical methods, Monte-Carlo simulation, Bayesian networks and Markov-Chain Monte-Carlo methods. The best method to provide the results required from the available data was assessed to be Monte-Carlo simulation. The simulations were implemented using the statistical programming language R, which was the package used for the data analysis. Methods of performing the simulations in Excel™ were also demonstrated.

Data on intakes of tap water by the population was available from drinking water consumption surveys (DWCS) in 1978, 1995 and 2008 (referred to here as DWCS1978 etc), a survey for the Food Standards Agency (FSA) in 2001 (NDNS 2001) and a survey of consumption by children under 16 for the DWI in 2011 (DW16 2011). The coverage and level of detail available varied between these. The DWCS 1978 and DWCS 1995 surveys included adults and children, whereas DWCS 2008 and NDNS 2001 included only adults. The only results available from the DWCS surveys were in the form the frequencies of different intakes. The results for individual intakes were available from NDNS 2001 and DW16 2011, which gave other information, such as the age and weight of the individuals. The reports of the DWCS analysed the effects of these factors, but did not give the data. All the surveys used 7-day diaries to record intake, but with two different methods of assessing the volume consumed. DWCS 1978 and NDNS 2001 used measured or weighed amounts, whereas the others relied on standard vessel volumes and estimates of the proportion filled and drunk.

The mean tap water intake by adults appeared to increase from 1.084 l/d in DWCS 1978 to 1.275 l/d in DWCS 1995 and to 1.294 l/d in DWCS 2008. The first increase was statistically significant, but the second was not. The mean intake in NDNS 2001 was 1.103 l/d, which was significantly different from the results for DWCS 1995 and 2008. It should be noted that the differences in intake coincide with the different methods used

to estimate the volumes used by the surveys, so it is possible that the apparent change in intake is an artefact of the method of estimation of consumption. The mean intake by all children in DW16 2011 was 0.551 l/d

Intake by adults and children tended to increase with both age and body weight, but the variability within each weight or age group was very high. The results for males and females in different surveys were inconsistent: in DWCS 1978, NDNS 2001 and DW16 2011, males drank more tap water than females; in the other surveys they drank less. It is interesting to note that the difference in the results for adults coincides with the difference in the recording method.

Probability distributions were fitted to the water intake data for each survey using maximum likelihood estimation. Four distributions were tested: (1) lognormal, (2) normal fitted to the square root of intake (as a continuous analogue of the Poisson distribution), here referred to as *normal (square root)*, (3) Weibull and (4) gamma. None gave a consistently better fit than the others; the gamma distribution generally gave good results and was selected for the simulation. Distributions were fitted to the complete populations and to the groups by sex and age or weight where they were available.

Data on the concentrations of chemicals in drinking water was available from compliance data collected by the water companies and supplied to the DWI. Data sets containing iron, lead, selenium, sodium and manganese concentrations from 2004 and 2010 were used in the study. Additional data sets on lead from some of the intervening years and less detailed data on lead from 1994 were used to look at the trends in exposure to lead over time. Some data on copper from 2010 was also included. Several distributions were fitted to the data using maximum likelihood estimation. In most cases, the best fit was obtained by the lognormal distribution, with the exception of selenium, for which the exponential distribution fitted better.

The data sets presented two problems: most sets had very long tails containing a few high values and several had a large proportion of small values that were not precisely determined, but recorded as below the limit of detection (LoD), for which a value was given. The problem of data below the LoD, known as left-censored data, was dealt with by using a version of maximum likelihood estimation that was designed for censored data. The long tails of the distributions could not be fitted by any standard distributions: all of them underestimated the frequency of rare extreme values, which are potentially important when assessing exposure probabilistically. As the data sets were large (over 10,000 values), the method chosen for the simulation was to sample from the original data instead of using the fitted distributions for most of the range. The distributions were used to generate substitute values for those lying below the LoD. The simpler alternatives of substituting 0, LoD/2 or the LoD were also tested.

Reference values for intakes of the substances being considered were taken from authoritative UK and other sources. There were three types of reference values. (1) Reference Nutrient Intakes (RNIs) for required minerals giving the amount that is sufficient for about 97% of the population (or similar measures). These existed for all the substances other than lead. (2) Safe limits for long-term intake, such as Acceptable

Daily Intake (ADI) or Provisional Maximum Tolerable Daily Intake (PMTDI), for potentially harmful substances. These were found for all the substances except manganese. (3) Intakes from sources other than tap water, such as dietary intake.

The simulation was written in R with a bespoke user interface and run with every intake distribution (i.e. the distributions fitted to the full sample in each survey and, where possible, the groups by sex and age or weight) and the 2010 concentration data. Additional runs were carried out using some of the intake distributions with the earlier sets of concentration data to explore changes over time. In each iteration of the simulation, a pair of values was sampled randomly from the appropriate water intake distribution and concentration data set, as described above, and multiplied to give a value for chemical intake. Sampling from the two distributions was independent: that is, no correlation between water intake and concentration was considered. Each run consisted of 100,000 iterations of the simulation, from which summary statistics were derived, including the mean, median, 99th and 99.9th percentile.

The results of the simulations have shown that exposure to metals in tap water is highly variable. The 99.9th percentile exposure can be up to 45 times the mean and 200 times the median. It should be emphasised that the percentiles relate to the chance of individual daily exposures, not long-term intake.

The method of substitution for values less than the LoD had only moderate effects on the estimation of the mean and percentiles up to the 75th for those substances with a high proportion of samples reported at the LoD (e.g. lead and iron), and smaller effects on the statistics of the other substances. The higher percentiles were unaffected in all cases. For simpler exposure assessments, substitution by either the LoD or LoD/2 would probably give acceptable accuracy.

Exposure to iron, lead, selenium and manganese predicted by the simulations appear to have decreased by about 40% between 2004 and 2010 due to falling concentrations in tap water. For lead this is part of a long term trend, having previously decreased by 40% between 1994 and 2004. In contrast, the exposure to sodium appears to have increased slightly.

Comparing the predicted exposures with Reference Nutrient Intakes (for required nutrients) and Acceptable Daily Intakes or other recommended maximum intakes, we found for adults, using 2010 concentration data, that:

- For iron, selenium, sodium and manganese, the 99.9th percentile exposures were much less than the RNIs. In each case, the RNI is much lower than the ADI or similar upper limit.
- For copper, the 99.9th percentile exposure slightly exceeded the RNI and the intake from other sources, but the mean was very much smaller than the RNI, so tap water may occasionally make a significant contribution to the requirement for copper. The 99.9th percentile exposure was less than 10% of the PMTDI.
- For lead, the 99.9th percentile exposure was about 40% of the ADI in the worst case and the mean exposure was about 1% of the ADI. The ADI is being

superseded by BMDL values. The mean exposure was about 5% of the $BMDL_{10}$ for nephrotoxity, which lay between the 99th and 99.9th percentiles of exposure in the worst case. The ADI and the BMDL both relate to lifetime exposure, not acute effects, so the mean exposure is the most appropriate comparison.

For children under 16, using 2010 concentration data, we found that:

- For iron, the 99.9th percentile of predicted exposure was less than 5% of the RNI in all cases. In the worst case, the 99.9th percentile was less than 5% of the PMTDI.
- For selenium, the 99.9th percentile of predicted exposure was less than 20% of the RNI in most cases and less than 50% in the case of the youngest group. The mean exposure was less than 4% of the RNI. Thus, tap water may occasionally, but not persistently, be a significant contributor to nutrient intake. The 99.9th percentile is 0.01% of the Upper Safe Level in the worst case.
- For sodium, the 99.9th percentile of predicted exposure for the youngest group (worst case) was about 30% of the RNI and the mean was about 3% of the RNI. The intake from other sources normally exceeds the RNI. Therefore, tap water may occasionally, but not frequently, be a significant contributor to sodium intake.
- For manganese, the 99.9th percentile of predicted exposure was less than 2% of the adequate daily intake for all groups aged 4 years and upwards. For the 0–3 years group, the 99.9th percentile was less than 4% of the adequate daily intake for children aged over 6 months, but 7 times the adequate daily intake for babies up to 6 months. The data set is insufficient to allow this age group to be simulated separately, but it is possible that the required nutrient intake may occasionally be exceeded for babies up to 6 months. However, manganese has low acute toxicity, and no PMTDI has been set.
- For lead, the 99.9th percentile exposure was less than half of ADI (for lifetime exposure) for most groups, but exceeded it slightly for the lightest group. The mean exposure was about 3% of the ADI in the worst case and about 1% of the ADI for the other groups. The mean exposure was less than 6% of the $BMDL_{01}$ (for long-term exposure) for developmental neurotoxicity in most cases and less than 20% in the worst case. The $BMDL_{01}$ was between the predicted 99th and 99.9th percentiles for most groups and between the 95th and 99th percentiles in the worst case. Thus the probability of persistently exceeding this level is relatively small.

Similar methods could be applied to other substances found in tap water, such as Trihalomethanes (by-products of chlorination). There are significant routes of exposure other than ingestion for these chemicals, notably skin contact and inhalation when bathing. A more complex model would therefore need to be constructed to adequately represent exposure to these substances.